PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/82, 5/10, A01H 5/00 — A1

(11) International Publication Number: WO 91/13993

(43) International Publication Date: 19 September 1991 (19.09.91)

(21) International Application Number: PCT/US91/00887

(22) International Filing Date: 14 February 1991 (14.02.91)

(30) Priority data:
489,175 5 March 1990 (05.03.90) US

(60) Parent Application or Grant
(63) Related by Continuation
US 489,175 (CON)
Filed on 5 March 1990 (05.03.90)

(71) Applicant *(for all designated States except US)*: THE UPJOHN COMPANY [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : SLIGHTOM, Jerry, Laverne [US/US]; CHEE, Paula, Po [US/US]; 3305 Lorraine Avenue, Kalamazoo, MI 49008 (US).

(74) Agent: THE UPJOHN COMPANY; Corporate Patents & Trademarks, Kalamazoo, MI 49001 (US).

(81) Designated States: AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL (European patent), NO, PL, RO, SD, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PROTEIN EXPRESSION VIA SEED SPECIFIC REGULATORY SEQUENCES

(57) Abstract

This invention relates to seed specific expression cassettes for regulation and expression of proteins important for pharmacological and industrial needs in plants. These cassettes may be rapidly constructed by joining regulatory elements via restriction enzyme sites synthesized into the amplification oligomers used to clone the regulatory sites. In one method of the invention the promoter region of the phaseolin gene is attached to an intronless gene coding for bovine somatotropin (Bst) and linked to a poly-A signal from either phaseolin, Bst, or both, for transfer into a plant genome.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

PROTEIN EXPRESSION VIA SEED SPECIFIC REGULATORY SEQUENCES

FIELD OF THE INVENTION

This invencion relates to seed specific expression casetces for use in regulation and expression of proteins in plants.

BACKGROUND OF THE INVENTION

In many plant species, seed storage proteins represent the predominate source of amino acids and nutrition for the germinating plant. It is well known that genes which code for storage proteins are strictly regulated. Seed storage proteins are produced in large quantities at a determined stage of development; expression is both tissue specific and developmentally controlled. These proteins are generally localized in specialized organelles called protein bodies or protein storage vacuoles and are essentially protease free. Because the seeds of many plants can readily synthesize large amounts of these storage proteins it is believed that the biological machinery within seed cells can be put to use synthesizing proteins of importance to pharmacological and industrial needs. Thus, characterization of the storage proteins as well as the ability to direct protein synthesis toward the protein bodies is of continuing experimental, commercial, and nutritional interest.

The storage proteins of the major grain crops have been particularly well studied. In the legumes the major seed storage proteins are the globulins, represented by two types: the nonglycosylated 11 S proteins and the glycosylated 7 S fraction. In the common bean (Phaseolus vulgaris) about 50% of the total seed protein is represented by a group of three polypeptides (the α, β, and γ subunits) collectively known as phaseolin. In soybean (Glycine max) the major storage protein in known as β-conglycinin and its subunits (α, α', and β) represent up to 80% of total seed protein in that genus.

In corn (Zea mays), a representative monocot, the protein zein constitutes about 80% of corn seed protein. While there are many zein-related genes responsible for zein production, one type, the β-zeins, are represented in very low copy number and yet account for 10% f total seed protein.

However, the se ds of many dicot and mon cot plants are defi-

cient in certain essential amino acids which require the addition of these amino acids when these seeds ar used as the predominate s urce of prot in in the human or animal diet. For a number of years investigators have attempted to couple the high level expression of seed storage proteins with recombinant engineering techniques to produce plants with enhanced nutritional value. For instance, while legume seed contain high levels of storage protein, the protein is deficient from a nutritional standpoint in that the globulins are relatively low in sulfur-rich amino acids, making legume seed a nutritionally incomplete foodstuff. Recombinant techniques could be used to increase the amount of sulfur in the seed storage proteins or direct synthesis of heterologous sulfur-rich proteins. The development of seed specific promoter cassettes for monocotyledons plant species is also important since the seeds from these plants are widely used as a food source for animals and humans. Thus, in a similar manner, zein, which is lacking in the essential amino acid lysine, can be made nutritionally complete.

On an industrial scale, such techniques can be used to direct heterologous proteins or enzymes into the storage vacuoles. As protease-free zones within the seed, such expression would serve as an abundant and easily purified source of protein. Thus, plant seed could provide an alternative means for producing hormones and enzymes needed for commercial and industrial use.

By using genetic engineering techniques it is already possible to transfer genes into many plant species; direct expression of homologous (native to the host) and heterologous (foreign) plant genes is well established. Numerous reports indicate that seed storage protein genes have been transferred to other plants with expression in both a tissue specific and developmentally regulated manner, Beachy, et al., EMBO 4:3047-3050, (1985), as well as into protein bodies. Greenwood and Chrispeels, Plant Physiol. 79:65-71 (1985).

The expression of non-plant genes in plants has also met with success, but on a more limited level. It has been suggested that the difficulty in attaining expression of non-plant proteins is due to a failure of the plant splicing machinery to recognize introns (also called intervening sequences) present in the non-native transcribed

pre-mRNA, K. Hartmuth and A. Barta, Nucleic Acids Res. 14: 7513-28 (1986), since removal of introns is necessary in order for the mRNA to be functional. See, M. Green, Annu. Rev. Genet. 20: 671-78 (1986). For instance, Barta, et al, Plant Molecular Biol. 6:347-357 (1986), report construction of a chimeric gene in which the complete human growth hormone (hGH) gene is flanked by DNA fragments coding for the promoter and polyadenylation sites of the nopaline synthase gene. When the chimeric gene is introduced into sunflower and tobacco callus tissue the gene is transcribed to the pre-mRNA stage. However, the transcript is non-functional in the plant tissue as the pre-mRNA is not further processed by the plant to mRNA. Thus the foreign protein, hGH, is not synthesized.

An alternative to direct expression of foreign proteins is to make a gene construct wherein the gene encoding for the desired protein is inserted within a nonfunctional region of a plant protein. European Patent Application 319 353 discloses a method for expressing a desired protein by modification of the hypervariable region of a determined storage protein by an insert encoding the desired peptide. In this way a modified storage protein is expressed which has contained within it amino acids which code for the foreign protein. The foreign protein is then cleaved by enzyme digestion, isolated and purified. However, the insert must be in reading phase with the non-modified part of the storage protein on either side of the insert, thus effectively limiting the size of the foreign protein which may be expressed to the length of the hypervariable region of the storage protein.

The present invention overcomes the problems outlined above by providing for seed specific expression of a desired protein comprising fusion of a gene with a seed specific gene regulatory sequence.

INFORMATION DISCLOSURE STATEMENT

The present invention uses a gene amplification technique which is a modification of the polymerase chain reaction (PCR) method described in U.S. Pat. 4,683,195.

European Patent Application 255 378 provides nucleotide sequences for seed specific transcriptional regulation. However, the reference does not teach the translation cassettes of the present invention.

European Patent Application 295 959 refers to DNA expression cassettes for transcription of the brazil nut sulfur-rich 2S seed storage protein. The reference neither teaches nor discloses the translational expression cassettes of the present invention.

European Patent Application 319 353 refers to a process for the production of useful biologically active polypeptides through the modification of plant genes encoding S2 albumins and is discussed above.

An East German Patent Application, DD 263-081-A, reports a process for the production of useful proteins via a vector containing a regulatory sequence isolated from a plant gene.

Barta, A., et al, "The Expression of Nopaline Synthase-Human Growth Hormone Chimaeric Gene in Transformed Tobacco and Sunflower Callus Tissue", Plant Molecular Biol, 6:347-357 (1986), report construction of a chimeric gene containing the complete human growth hormone (hGH) gene flanked by the promoter and polyadenylation site of the nopaline synthase gene and is discussed above.

Beachy, et al, "Accumulation and Assembly of Soybean β-Conglycinin in Seeds of Transformed Petunia Plants", EMBO, 4:3047-3053 (1985), report the accumulation of the α'-subunit of the soybean seed storage protein β-conglycinin in the seeds of transformed petunia cells. The reference does not teach or disclose the cassettes of the present invention.

P.P. Chee, et al, "Expression of a Bean Storage Protein 'Phaseolin Minigene' in Foreign Plant Tissues", Gene, 41:47-57 (1986), report that stable and functional phaseolin mRNA can be produced in tobacco by the use of an mutant phaseolin gene lacking the native introns.

The gene coding for the β-zein 15Kd protein and regulatory regions are used in this invention. The sequence and isolation of this sulfur-rich zein protein is disclosed in Pedersen, et al, "Sequence Analysis and Characterization of a Maize Gene Encoding a High-Sulfur Zein Protein of M_r 15,000", J. Biol. Chem. 261:6279-6284 (1986).

Sengupta-Gopalan, et al, "Developmentally Regulated Expression of the Bean β-Phaseolin Gene in Tobacco Seed", Proc. Natl. Acad. USA, 82:3320-3324 (1985), report the transformation of tobacco with a 3.8-

kB phaseolin DNA fragment containing the coding region and flanking DNA necessary for c rr ctly regulated expression of the gene. However, the present invention uses specific linkage sequences not taught by the reference.

The complete nucleotide sequence of the French bean seed storage protein phaseolin is disclosed in Slightom, et al, "Complete Nucleotide Sequence of French Bean Storage Protein Gene: Phaseolin", Proc. Natl. Acad. USA, 80:1897-1901 (1983). Some of the methods reported therein are used in the present invention.

SUMMARY OF THE INVENTION

The present invention provides a seed specific expression cassette which has a promoter derived from either phaseolin, α'-subunit of β-conglycinin, or β-zein 15Kd; a translation initiation signal from either phaseolin, the α'-subunit of β-conglycinin, β-zein 15Kd, or an animal gene; a gene derived from brazil nut sulfur-rich seed storage protein or an animal gene; a translation termination region derived from either the animal gene, phaseolin, or β-zein 15Kd; and at least one polyadenylation region from phaseolin, the animal gene, or β-zein 15Kd; wherein the regulatory sequences are operably linked to one another in such a manner that the gene is expressed in seed or seed storage protein bodies. The invention further provides enzyme restriction sites incorporated into the sequences which link the regulatory elements and genes. The invention alao provides that certain of the regulatory elements be derived from the same genetic source.

More particularly, the present invention provides an expression cassette in which the promoter is from phaseolin or α'-subunit of β-conglycinin, the animal gene and remaining elements are from bovine somato-tropin, and the restriction sites contained within the linkers are NcoI and HindIII.

More particularly, the present invention provides an expression cassette in which the promoter is from phaseolin or α'-subunit of β-conglycinin, one poly-A signal is from phaseolin, the remaining regions, including a second poly-A site, are from bovine somatotropin and the restriction sites are NcoI and HindIII.

Most particularly, the present invention provides an expression cassette for the expression and regulation of bovine somatotropin in

seed storage bodies.

In another embodiment of the invention the gene is from brazil nut 2S storage protein or subunits thereof, the promoter and translation initiation signal are derived from the same genetic source, and the translation termination signal and poly-A region are from phaseolin or β-zein 15Kd.

The invention also provides for cells, plants, and seeds containing the expression cassettes disclosed here.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a plant in which a gene coding for a specified protein is incorporated into the plant's genome in a manner such that regulation and expression of the specified protein is under the control of a seed specific regulatory sequence. The protein which is of interest and for which expression is desired may be a plant protein homologous or heterologous to the host plant, or on the other hand, may be normally produced in a virus, microorganism, or other animal tissue.

The following examples illustrate specific embodiments of the invention. We have used a modified form of the polymerase chain reaction technique to combine various seed storage protein gene regulatory elements (promoters, translation initiation sites, translation termination sites, and transcription polyadenylation (poly-A) signals) with protein coding regions to form expression cassettes which direct and control the production of various modified proteins in the seeds of plants. These expression cassettes may then be used to transform plant cells from which whole plants are derived. Gene regulatory elements are used which may be inserted into hosts of either monocot or dicot plant species. Of particular interest as host plants are those which are cultivated for high levels of seed production. Such hosts include wheat, corn, rice, etc. (monocots) and soybean, common bean, pea, etc. (dicots). Corn and soybean represent ideal hosts and seed storage protein regulatory regions isolated from *Phaseolus vulgaris* (phaseolin gene), *Glycine max* (promoter from the α'-subunit of β-conglycinin gene) and *Zea mays* (zein 15 Kd peptide gene) are preferred as sources of regulatory regions. Proteins which would have value when expressed in plant seeds includ those designed for nutritional improvement (e.g., high

sulfur soybean and common bean, high lysine corn, etc.), pharmacological importance (enzymes and hormones),and industrial value (enzymes). We have demonstrat d seed specific expression cassettes containing genes designed to improve the nutritional value of seed protein meal by encoding a high-sulfur gene. W have also demonstrated the production and segregation of bovine somatatropin (Bst) protein into protein storage vacuoles.

Numerous cloning vectors are available which have an appropriate restriction site, a replication system, and a marker. As a cloning vector we have selected pUC18, Norrander et al, Cene (1983)26:101-106, containing an ampicillin resistance marker. Other suitable vectors are well known and may be easily adapted for use with the specific restriction sites of the present expression cassettes.

For the genetic engineering experiments described in the examples one can use many different methods of obtaining the desired sequence junctions between regulatory regions and coding DNA, such as, site specific mutagenesis, linkers, and/or polymerase chain reactions (PCR). We have selected to use a modification of PCR, which we call custom polymerase chain reaction (CPCR) as the technique for the addition of regulatory regions with coding DNA. CPCR allows one to incorporate specific enzyme restriction sites directly into the linkage oligomers as they are synthesized. The oligomers are then used in PCR amplification of a desired gene or regulatory element. The advantage to this modification is a more rapid construction of expression cassettes from the amplified sequences.

We include in the following examples expression cassettes for transcriptional and translational fusion; two dicot specific seed promoter cassettes; and one monocot specific seed promoter cassette. Those skilled in the art will recognize that these cassettes may be utilized in a variety of hosts. In addition, the incorporation of these cassettes into a host plant genome may be accomplished by a variety of conventional methods. Such methods include transformation with Ti-DNA (A. tumefaciens, A. rhizogenes, etc.), protoplast fusion, injection, electroporation, and microprojectile bombardment. These plant seed expression cassettes are used to express a high sulfur seed storage protein gene in the seeds of both dicot and monocot plants and to express a non-plant gene, bovine somatotropin (Bst), in

the seeds of dicot plants.

Transgenic plants containing the heterologous gene of interest are obtained through the R1 generation. Seeds from these plants, R2 seeds, are stained with protein antibody for TEM visualization of the heterologous protein localized within the protein storage bodies..

The following abbreviations are used throughout this disclosure to denote gene regulatory elements:

P = Promoter

SP = Translation Initiation Signal or Signal Peptide

TT = Translation Termination Signal

S = Polyadenylation (Poly-A) Signal

The following abbreviations are used throughout this disclosure to indicate source of the gene regulatory elements:

bv = genus *Bovine*

bz = Bertholletia excelsia (Brazil nut)

gm = *Glycine max*

pv = *Phaseolus vulgaris*

zm = *Zea mays*

The following abbreviations are used throughout this disclosure to indicate proteins:

"Bst" means bovine somatotropin protein. "BST" means the gene encoding Bst.

"Bz" means a sulfur-rich protein isolated from Brazil nut seed. "Bz-1", Bz-2", and "Bz-3" are subunits of Bz. These proteins are described in further detail below. "BZ", "BZ-1", "BZ-2", and "BZ-3" are the genes encoding the respective proteins.

The following terms are used throughout this disclosure and are defined as follows:

"Seed specific expression cassette" means a DNA construct containing gene regulatory elements and a gene encoding a protein of interest, in a suitable vector for incorporation into a plant genome, such that the expression of the gene is under the control of the gene regulatory regions, the regions being of the specificity, orientation, and reading frame such that expression of the gene is in the plant seed.

"Expression cassette" means a DNA construct lacking at least a gene encoding a protein of interest.

"Oligomer primer" means a nucleotide sequence needed for PCR amplification of a gene or gene r gulatory element. The primers used in this invention have been synthesized to incorporate specific restriction enzyme sites into the primer.

Charts 1 through 22 illustrate the constructions of the present invention. The following conventions are used to illustrate plasmids and DNA fragments:

(1) The single line figures represent both circular and linear double-stranded DNA.

(2) Asterisks (*) indicate that the molecule represented is circular. Lack of an asterisk indicates the molecule is linear.

(3) Endonuclease restriction sites are indicated above the line.

(4) Genes are indicated below the line.

(5) Distance between genes and restriction sites are not to scale. The charts show their respective positions only.

The methods used here employ many standard recombinant DNA procedures as well as those for creating transgenic plants and are well known by those skilled in the art. Standard methods and laboratory techniques for recombinant experimentation can be found in T. Maniatis et al., "Molecular Cloning; A Laboratory Manual", Cold Spring Harbor Laboratory (1982) (recombinant DNA experimental protocols and methods); Miller, "Experiments in Molecular Genetics", Cold Spring Harbor Laboratory (1972) (culture techniques); Southern, J. Mol. Biol. 98: 503 (1975) (clone verification); Sanger, et al, Proc. Natl. Acad. Sci. USA 74:5463 (1977) (DNA sequencing). S.B. Gelvin and R.A. Schilperoot, "Plant Molecular Biology" (1988), supplies techniques for use in transforming and cloning plants.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

EXAMPLE 1

CONSTRUCTION OF AN EXPRESSION CASSETTE CONTAINING THE PHASEOLIN PROMOTER (pvP) AND POLYADENYLATION SIGNAL (pvS) FOR TRANSCRIPTION FUSION TO CODING DNA

The major seed storage protein gene of *Phaseolus vulgaris* (common bean) is known as phaseolin. The complete sequence, including 5' and 3' regulatory elements, and cDNA counterpart have been published by Slightom et al. (1983). As a source of phaseolin regulatory regions we use the clone pPhasmin which has the intronless

cDNA counterpart in place of the g nomic coding and intron regions, as described by Cramer et al., Pr c. Natl. Acad. USA, 82:334-338 (1985). Oligomer primers for amplification of phaseolin promoter (pvP) are constructed to contain HindIII (in the 5' primer) and NcoI (in the 3' primer) sites. The *Nco*I site is designed to include a translation initiation site which conforms to the rules described by Kozak (1986) by containing the sequence AAXXATCG where X represents G, A, T or C. Sequences of the oligomers used for the amplification and engineering of the phaseolin promoter are shown in Chart 1(a). Using CPCR technology the restriction sites are fused to the 500 b.p. phaseolin promoter region resulting in the fragment designated pvP in Chart 1(b).

In a similar manner oligomer primers are synthesized to flank the phaseolin polyadenylation (poly-A) signal. In this case, however, the 5' primer incorporates a NcoI site and the 3' primer a HindIII site. The nucleotide sequences of these two primers is shown in Chart 1(c). Amplification results in a 1100 b.p. fragment designated pvS in Chart 1(d). The pvP and pvS fragments were subjected to digestion by *Nco*I followed by ligation of the fragments (Chart 1(e)). The ligated pvP + pvS fragment (1600 b.p.) is subjected to *Hin*dIII digestion, purified on polyacrylamide gel, and cloned into *Hin*dIII digested pUC18. The clone containing the correct pvP + pvS sequences (confirmed by restriction enzyme mapping and nucleotide sequencing) is referred to as pUC18pvPpvS in Chart 1(f). This expression cassette is used with coding regions which either contain their own translation initiation signal or do not need a signal for proper targeting of the protein product. in addition, any coding regions inserted into this cassette must include a translation termination signal.

EXAMPLE 2

CONSTRUCTION OF A PHASEOLIN EXPRESSION CASSETTE CONTAINING PHASEOLIN PROMOTOR (pvP), SIGNAL PEPTIDE (pvSP), TRANSLATION TERMINATOR (pvTT), AND POLY-A SIGNAL (pvS) FOR TRANSLATIONAL FUSIONS

In this example oligonucleotide primers are designed for the amplification of promoter and terminator fragments from pPhasmin which contain the phaseolin signal peptide (SP) and translational termination (TT) signals, respectively. N-terminal analyses of

phaseolin protein by Parran et al., Phytochemistry 26: 335 (1987), show that cleavage is ragged and occurs after amino acid residue 23 and before amino acid 27. Thus, to ensure that the cassette contains the amino acids necessary for signal peptide cleavage the cassette is designed to include through residue 31. For CPCR amplification of the pvPpvSP fragment oligomers are designed which contain an inframe 5' *Hin*dIII site and an inframe 3' *Nco*I site; these sequences are shown in Chart 2(a). Amplification yields the 580 b.p. fragment designated pvPpvSP in Chart 2(b).

To ensure complete reading of the translation termination sequence, the 4 amino acids 5' of the translation termination codon are retained. These phaseolin 3' regulatory regions are obtained using the primers shown in Chart 2(c). Again, these primers contain an inframe *Nco* I site in the 5' oligomer and a *Hin*dIII site in the 3' oligomer. Amplification yields a 1200 b.p. fragment referred to as pvTTpvS in chart 2(d).

Both fragments, pvPpvSP and pvTTpvS, are subjected to digestion by *Nco*I and the exposed *Nco*I sites are ligated together to yield the fragment referred to as pvPpvSPpvTTpvS in chart 2(e). This fragment is then digested with *Hin*dIII to expose CPCR engineered 5' and 3' sites followed by cloning into *Hin*dIII digested pUC18. The phaseolin expression cassette referred to as pUC18pvPpvSPpvTTpvS in chart 2(f) is obtained and checked by restriction enzyme mapping and nucleotide sequencing to determine that it is correct.

This seed specific expression cassette is used with coding regions which need a seed specific signal peptide and an inframe translation termination region.

EXAMPLE 3

CONSTRUCTION OF AN EXPRESSION CASSETTE CONTAININC THE PROMOTER RECION OF THE CLYCINE MAX SEED STORAGE PROTEIN α'-SUBUNIT OF β-CONGLYCININ (gmP) AND PHASEOLIN POLY-A SICNAL (pvS)

In soybean (Clycine sp.) β-conglycinin represents the major seed storage protein. The level of expression is very high; each of the two to four subunits account for between 12 to 20% of the total protein in the seed. The nucleotide sequence of the promoter region of *Glycine max* α'-subunit of β-conglycinin has been published by Doyle et al., J. Biol. Chem. 261:9228 (1986).

The promoter (gmP) is amplified from 250 ng t tal *Glycine max* DNA and, by using CPCR, a *Hin*dIII site is added t the 5' end as well as an *Nco*I site at the translation initiator codon (ATC). Using the published sequence oligomers are designed for amplification of the promoter (oligomer sequences are shown in Chart 3(a)). The sequence of the amplified 600 b.p. gene, designated gmP in Chart 3(b), is verified by restriction enzyme mapping and nucleotide sequencing for accuracy.

The fragment gmP was subjected to digestion by *Nco*I and ligated to the *Nco*I digested fragment pvS (described in Example 1) to obtain the fragment referred to as gmPpvS in Chart 3(c). The resulting fragment, gmPpvS, is digested with *Hin*dIII, ligated into *Hin*dIII digested pUC18, and the clone pUC18gmPpvS is obtained. (See Chart 3(d)). This expression cassette is used to direct seed specific expression of coding regions which require transcriptional fusions, similar to that shown in Example 1. However, when incorporated into the Glycine genome the level of expression may be somewhat higher than the Phaseolus expression cassette of Example 1, due to the more native environment for the Glycine cassette.

EXAMPLE 4

CONSTRUCTION OF AN EXPRESSION CASSETTE CONTAINING THE *GLYCINE MAX* α'-SUBUNIT OF β-CONGLYCININ PROMOTER (gmP) AND SIGNAL PEPTIDE (gmSP) REGIONS AND THE PHASEOLIN TRANSLATION TERMINATION (pvTT) SIGNAL AND POLYADENYLATION (pvS) SIGNALS.

In this example we describe construction of an expression cassette and plasmid which is used for the expression of modified seed storage protein genes or heterologous genes in the seeds of dicotyledonous plants. Oligomers are used to amplify the promoter and signal peptide region from *Glycine max* genomic DNA. For the oligomer which is read 5' → 3' the same 5' primer as described in Example 3 is used. To ensure complete reading of the signal peptide, the 3' oligomer primer is designed to hybridize just beyond the signal peptide cleavage point described by Doyle et al, (1985), which occurs between amino acid residues 25 and 26. Therefore, we have selected to add the *Nco*I site after amino acid residue 29. The primers used for this amplification are shown in Chart 4(a).

The result of this amplification, referred to as gmPgmSP in

Chart 4(b), is a fragment of 690 b.p. Authenticity of this fragment is checked by nucleotide sequencing. It is then subjected to digestion by *Nco*I and ligated to the *Nco*I cut fragment pvTTpvS described in Example 2. The resulting fragment, referred to as gmPgmSPpvTTpvS in Chart 4(c), was subjected to digestion by *Hin*dIII and cloned into *Hin*dIII digested pUC18. The clone (designated pUC18gmPgmSPpvTTpvS in Chart 4(d)) is isolated and its structure checked by restriction enzyme mapping and nucleotide sequencing. This expression cassette is used for the expression of modified seed storage protein genes or foreign genes in the seeds of dicotyledonous plants.

EXAMPLE 5

CONSTRUCTION OF AN EXPRESSION CASSETTE CONTAININC THE PROMOTER (zmP) AND POLYADENYLATION SIGNAL (zmS) FROM THE ZEA MAYS 15 kDAL SEED STORAGE PROTEIN GENE.

The zein proteins constitute about 80% of the total seed protein of *Zea mays* (corn); however, there are many zein-related genes responsible for the production of these zein proteins. Recently, one particular member of the zein gene families was found to be represented in low copy number, but responsible for the production of about 10% of the total zein proteins. These genes are referred as β-zeins (or 15 Kd) genes and the sequence of one member has been described by Pedersen et al. (1986). However, the nucleotide sequence of the β-zein gene reported by Pedersen only extends about 233 b.p. 5' of the translation initiation, thus the complete β-zein gene promoter may not be present. To obtain more of the β-zein promoter we isolated a second gene, referred to as pZein15K, and our nucleotide sequence analysis (Chart 5) shows that this clone contains an additional 110 b.p. of 5' flanking DNA, giving a total promoter region of about 350 b.p. to the translation initiation codon (see Chart 5). Analysis of many zein type genes has shown that sequences responsible for controlling the developmental expression of seed specific genes are located within these 350 b.p. (see review by Thompson and Larkins, BioEssays, 10:109-113 1989).

Clone pZein15K (Chart 6(a)) provides the starting material for the CPCR amplification of its promoter and polyadenylation regions. However, prior to amplification an indigenous *Nco*I site in the

promoter region was removed by flushing with mung bean nucleas following the procedure of Maniatis, et al, (1982), which caused a loss of four base pairs (CATG). These four base pairs are located between the CCAAAT and TATAA elements and thus should not affect the function of the promoter. This modified β-zein gene promoter region was amplified using CPCR to add a *SalI* site and a *NcoI* site to the 5' and 3' primers, respectively. The sequences of the oligomers used to accomplish this are shown in Chart 6(b). The amplified product was a 350 b.p. fragment which is referred to as zmP in Chart 6(c). Additional CPCR oligomer primers were used to amplify the polyadenylation region of pZein 15K so as to add a (5') *NcoI* site and a (3') *SalI* site. Oligomers sequences are shown in Chart 6(d). The amplified product was a fragment of approximately 700 b.p. which is referred to as zmS in Chart 6(e). The two amplified fragments, zmP and zmS, are subjected to digestion with *NcoI* followed by ligation and isolation of the fragment referred to as zmPzmS in Chart 6(f). zmPzmS is digested with *SalI* and cloned into *SalI* digested pUC18. Clone pUC18zmPzmS (Chart 6(g)) is isolated and its structure verified using restriction enzyme mapping and nucleotide sequencing. The expression cassette pUC18zmPzmS is used for seed specific expression of modified seed storage protein genes or other genes which do not require a zein signal peptide and/or a zein translation terminator.

EXAMPLE 6

CONSTRUCTION OF A MONOCOT SEED SPECIFIC EXPRESSION CASSETTE CONTAINING THE β-ZEIN PROMOTER (zmP), SIGNAL PEPTIDE (zmSP), TRANSLATION TERMINATOR (zmTT) AND POLYADENYLATION SIGNAL (zmS).

This expression cassette is similar to that in Example 5, with the addition of the β-zein signal peptide and translation terminator signals. As in Example 5, CPCR adds a 5' *SalI* site and a 3' inframe *NcoI* site. Since in this case, however, the signal peptide cleavage point is located between amino acid residues 20 and 21, we have included amino acid sequences up to and including residue 25. Thus, the NcoI site is placed 3' of this residue. The oligomer primers used to engineer and amplify this zein promoter and signal peptide fragment are shown in Chart 7(a). This amplification results in a 400 b.p. fragment referred to as zmPzmSP in chart 7(b).

In a similar manner, CPCR amplification of the β-zein transla-

tion terminator and polyadenylation signal regions of pZein 15K is done using a linker engineered to contain a 5' inframe *NcoI site and* a 3' *SalI* site; sequences are shown in Chart 7(c). Amplification results in a 720 b.p. fragment referred to as zmTTzmS in Chart 7(d). The fragments zmPzmSP and zmTTzmS are subjected to digestion by *NcoI* and ligated together to obtain the fragment referred to zmPzmSPzmTTzmS in Chart 7(e). This fragment is then subjected to digestion by *SalI* and cloned into *SalI* digested pUC18. The resulting clone, pUC18zmPzmSPzmTTzmS (Chart 7(f)), is confirmed by restriction enzyme site mapping and nucleotide sequencing. This seed specific expression cassette pUCzmPzmSpzmTTzmS is used to regulate the expression of modified seed storage protein genes or other genes which may require a monocot or β-zein signal peptide and inframe translation terminator signal.

EXAMPLE 7

PHASEOLUS VULGARIS, GLYCINE MAX, AND ZEA MAYS SEED SPECIFIC EXPRESSION CASSETTES FOR THE EXPRESSION OF BRAZIL NUT 2S SULFUR-RICH SEED STORAGE PROTEIN GENE.

This example illustrates the use of the seed specific expression cassettes designed in Examples 1, 3, and 5 for the expression of a sulfur-rich (26% methionine + cystine) seed storage protein isolated from Brazil nut. The isolation and nucleotide sequence of this Brazil nut gene has been reported by Altenbach et al., Plant Mol. Biol. 8:239-250 (1987), and the cDNA clone referred to as pHS-3 contains all but 25 b.p. of the 5'-untranslated region, all of the coding DNA and the complete 3'-untranslated region, as illustrated in Chart 8. Because the gene encodes a protein which is normally found in a plant seed environment it is believed that its stable expression and proper transportation to seed storage protein bodied are characteristics which are inherent to its structure. However, an analysis of the sequence reveals a second *NcoI* site located near the translation termination codon. Thus for the ease of engineering we eliminated this *NcoI* site by using CPCR to replace the native NcoI site with an inframe *NcoI* site. This results in the second G being replaced with a C nucleotide and removes the internal *NcoI* site. The *NcoI* site added via CPCR is used to clone this region into an expression cassette which has an inframe translation termination codon. For

this construction w use tha promoter regions dascribed in Examples 1, 3, and 5 and translation termination and polyadenylation signals describad in Examples 2, 4, and 6.

CPCR amplification of the 8razil nut gene added was used to engineer flanking *Nco*I sites; the 5' *Nco*I site is added to take advantage of the naturally occurring sites located at the Brazil nut translational initiation codon (ATG), and the second *Nco*I site is removed and replaced by a *Nco*I site as described above. A further advantage of this modification is the addition of the amino acid tryptophan to the gene. Tryptophane is also deficient in the proteins of many plant seeds thus its addition results in a more nutritionally complete seed protein. The primers for the amplification of the Brazil nut coding region are shown in Chart 9(a). Amplification yields a 420 b.p. fragment which is referred to as BZ in Chart 9(b). This fragment is subjected to digestion with *Nco*I and inserted into the expression cassettes described earlier. The following clones are isolated, pUC18pvP*BZ*pvTTpvS, pUC18gmP*BZ*pvTTpvS, and pUC18zmP*BZ*zmTTzmS, and their structures are checked using restriction enzyme mapping and nucleotide sequencing. The structure of these clones are shown in Charts 9(c)(i), (ii), and (iii). These cassettes may then be removed and placed into an appropriate vector for transfer into plants. These transfer vectors may be used in either the *Agrobacterium* systems or in systems which utilize microprojectile bombardment of plant tissues.

EXAMPLE 8

USE OF *PHASEOLUS VULGARIS*, *GLYCINE MAX*, AND *ZEA MAYS* SEED SPECIFIC EXPRESSION CASSETTES FOR THE EXPRESSION OF A SYNTHETIC SULFUR-RICH SEED STORAGE PROTEIN GENE.

Because seeds from plants are deficient in more than one amino acid and because no single seed storage protein gene has been found which satisfies all of these needs, it is most likely that these needs will be accomplished via a synthetic coding region. To illustrate this possibility three oligomers are synthesized (shown in Chart 10) representing portions of the Brazil nut coding DNA. As shown in Chart 10(a), the first oligomer encodes the major portion of the 9 kD subunit (BZ-1), which contains 35% sulfur amino acids for the 48 amino acid peptide. Chart 10(b) shows a second oligomer con-

sisting of the complete 9 kD subunit (BZ-2) which encodes 70 amino acids containing 28.5% sulfur amino acids. Th third oligomer, Chart 10(c), is the complete 12 kD subunit (BZ-3) which encodes 93 amino acids containing 28% sulfur amino acids. BZ-1 was synthesized as a single 144 base oligomer, BZ-2 as a 210 base oligomer and BZ-3 as a 279 base oligomer. These oligomers were synthesized by utilizing the published sequence for the BZ gene and were completed on an Applied Biosystems, Inc. Model 380A DNA Synthesizer. The brazilnut gene fragments were complemented, purified, and had *Nco*I sites added on both 5' and 3' ends using CPCR amplification. As is true of the BZ gene in Example 7, each inframe *Nco*I site yields the codon TGG which encodes tryptophan, making resulting plant seed a more nutritionally complete foodstuff.

The primers used for amplification of BZ-1 are shown in Chart 11(a), those for BZ-2 in Chart 11(b), and those for BZ-3 in Chart 11(c). Because of the addition of the extra nucleotides the fragments isolated after amplification are about 15 b.p. larger than the original synthesized fragments BZ-1, BZ-2, and BZ-3. These fragments are subjected to digestion with *Nco*I and cloned into the *Nco*I cut expression cassettes described in Examples 2, 4 and 6. The following clones are isolated, pUC18pvPpvSPBZ-*1*, -*2*, *and* -3pvTTpvS (Charts 12(a)-(c)), pUG18gmPgmSPBZ-*1*, -*2*, *and* -3pvTTpvS (Charts 13(a)-(c)), pUC18zmPzmSPBZ-*1*, -*2*, *and* -3pvTTpvS (Charts 14(a)-(c)). The structure of these clones is checked by restriction enzyme mapping and nucleotide sequencing. The expression cassettes contained within these plasmids may removed and transferred into vectors used for *Agrobacterium* systems or may used with microprojectile bombardment of plant tissues.

EXAMPLE 9

SEED SPECIFIC EXPRESSION CASSETTES FOR THE EXPRESSION OF A FOREIGN GENE.

In this example we have selected to use the seed expression cassettes described above for the engineering of vectors capable of expressing the bovine somatotropin gene (Bst); this gene is also referred to as the bovine growth hormone (Bgh) gene. The sequence of the Bst gene has been determined by Woychik et al., Nucleic Acids Research, 10:7197-7210 (1982), however, this sequence is not accep-

table for expressi n in a plant system due to the presence of 4 introns in the native gene (Barta et al. 1986). To avoid this problem we use an intronless (minigene) construction of the Bst gene obtained from Professor Fritz Rottman, Department of Biology and Microbiology, Case Western Reserve University, Cleveland OH. This clone is referred to as pSVGH-1 (delta 1VS) and is used to provide the Bst coding region for CPCR engineering. As is clear from the prior examples, the use of CPCR allows the coding region to be engineered in such a manner to be accepted by any of the expression cassettes described in Examples 1 to 6. These arrangements can include the use of different gene regulatory regions which can be provided by the expression cassette or even from the Bst gene.

Bst protein is encoded in a gene which contains a 26 amino acid signal peptide followed by a coding region containing 191 amino acids. The clone pSVGH-1 also contains the indigenous Bst gene polyadenylation signal. For the Bst protein to be active its signal peptide must be removed, thus by using CPCR in combination with the cassettes described above we have the option of expressing it with or without its signal peptide, with or without a seed protein signal peptide, and with or without its polyadenylation signal. Thus many different construction can be derived for the expression of Bst in plant seeds. This example illustrates several of these expression cassettes coding for Bst.

The first series of constructions uses the seed storage protein gene expression promoters, described in Examples 1, 3 and 5, to construct transcriptional fusions genes which use the Bst signal peptide and polyadenylation signal. Using the clone pSVGH (delta 1VS) as starting material, CPCR is used to add an *Nco*l site to its 5' end (retaining the translation initiation signal) and adding a *Hin*dIII site 3' of the Bst polyadenylation signal. The oligomer primers of this construction are shown in Chart 15(a) and are used to amplify a 840 b.p. fragment referred to as bvSP*BST*bvTTbvS. This fragment is subjected to digestion by *Nco*I and cloned into the seed promoter fragments described in Examples 1 and 3, see Charts 15(b) and 16(a). These ligated seed promoter-Bst fragments are then digested with *Hin*dIII followed by ligation to *Hin*dIII digested pUC18 and the following clones isolated, pUC18pvPbvSP*BST*bvTTbvS and

pUC18gmPbvSP*BST*bvTTbvS, as shown in Charts 15(c) and 16(b), respectively. The correct sequence is verified using restriction enzyme site mapping and nucleotide sequencing. The inserts from these clones can be removed and transferred into either an *Agrobacterium* vector system or into a vector to be used for microprojectile bombardment of plant tissues.

To avoid the possibility of having an unstable mRNA product, the Bst gene can be cloned directly into the *Nco*I site located within the seed expression cassettes shown in Examples 1 and 3. To accomplish this only the 3' oligomer primer needs to be changed to include *Nco*I. This is done with the use of the oligomer primer shown in Chart 17(a) and CPCR engineering. The fragment containing the Bst gene obtained from this amplification is also about 840 b.p. in length, and is referred to as bvSP*BST*bvTTbvS in Chart 17(b). This fragment is digested with *Nco*I and cloned into the *Nco*I site of the plant expression cassette shown in Example 1. From this ligation, a clone designated pUC18pvPbvSP*BST*bvTTbvSpvS is obtained and is shown in Charts 17(c). In a similar manner, the expression cassette of Example 3 is inserted into bvSP*BST*bvTTbvS (Chart 18(a)) resulting in the plasmid pUC18gmPbvSP*BST*bvTTbvSpvS of Chart 18(b).

Additionally, by the use of another 3' oligomer primer a similar construction is derived which does not include the Bst polyadenylation signal. The 3' primer used to obtain the Bst such a fragment without its polyadenylation signal is shown in Chart 19(a). Amplification yields a fragment of about 650 b.p., referred to as pvSPBstbvTT in Chart 19(b). This is subjected to digestion by *Nco*I and cloned into the plant expression vector shown in Example 1, resulting in clone pUC18pvPbvSP*BST*bvTTpvS (Chart 19(c)). In addition, this Bst fragment (Chart 20(a)) is ligated into the vector of Example 3 which results in clone pUC18gmPbvSP*BST*bvTTpvS, shown in Chart 20(b).

EXAMPLE 10

SEED SPECIFIC EXPRESSION CASSETTES FOR THE EXPRESSION OF A FOREICN CENE.

In its active form, Bst does not include its 19 amino acid signal peptide, which is cleaved soon after synthesis (Woychik et al., 1982). However, in a plant system it is doubtful that this

signal peptide would be removed, thus its removal would have to be done as part of a purification procedure. In addition, the above constructions do not include the normal signal peptides for the plant seed storage protein genes to ensure correct targeting of the Bst peptide. The following constructions illustrate a translational fusion, made with the Bst protein gene, which does not include the Bst signal peptide but does contain the signal peptides from the plant seed storage protein genes. This is accomplished by using CPCR to amplify the part of the Bst gene which does not include the Bst signal peptide and cloning it into the expression cassettes described in Examples 2 and 4. The polypeptide resulting from expression from these fused gene is expected to initially include the seed storage signal peptides, which should be removed as a result of normal proteolytic cleavage to yield a Bst polypeptide which would be identical to the naturally occurring peptide except for the presence of two to four additional amino acids at its N-terminus. These additional amino acids are needed to ensure that the signal peptide cleavage site is recognized (see Examples 2, 4 and 6).

The appropriate Bst gene fragment can be obtained using CPCR amplification using the two primers shown in Chart 21(a) which amplifies a 600 b.p. fragment referred to as BSTbvTT in Charts 21(b) and 22(a). This fragment is subjected to digestion by *Nco*I and then ligated into the seed storage protein gene cassettes described in Examples 2 and 4 to obtain clones, pUC18pvPpvSP*BST*bvTTpvS (Chart 21(c)) and pUC18gmPgmSP*BST*bvTTpvS (Chart 22(b)). The engineered expression cassettes contained within these clones can be removed by digestion with *Hin*dIII followed by transfer to vectors designed to be used with *Agrobacterium* or microprojectile gene transfer systems.

EXAMPLE 11

PLANT CONTAINING A CASSETTE FOR EXPRESSION OF BST

A plant was produced which contained an expression cassette coding for the bovine somatotropin gene, BST. We began by removing the regulatory region of the phaseolin gene from clone pPhas 8.8 BglII, J.L.Slightom, et al, Proc. Natl. Acad. Sci. U.S.A. 80:1897-1901 (1983), by restriction endonuclease digestion with BglII and the fragment purified following standardized techniques. Next, we used the transformed SV40 plasmid, pSVbGH-1, (described earlier) contain-

ing a complementary DNA (cDNA) of bovine somatatropin (Bst) and a Bst polyadenylation region; these are fused downstr am of the SV40 promoter. Plasmid pSVbGH-1 contains a single restriction site for BamHI which lies between the SV40 promoter and the Bst gene. Finally, pPhasBst is made by digestion of pSVbGH-1 with BamHI followed by subsequent incorporation of the purified promoter region at the restriction site.

The region of the pPhasBst plasmid containing the Bst genes and the phaseolin regulatory region is removed by digestion with KpnI and inserted at the KpnI site of Agrobacterium binary vector pGA482G to maka plasmid pGA482G/PhasBst. pGA482G/PhasBst is used to transform Agrobacterium strain G58Z707. A.G. Hepburn, et al, J. Gen. Microbiol. 131:2961-2969 (1985). Finally, the resultant Agrobacterium is used to infect cultured tobacco leaf pieces, from which whole plants were derived.

Transgenic tobacco plants containing the PhasBst gene are obtained through the R1 generation. Saeds from theses plants, R2 seeds, are stained with Bst antibody for TEM visualization. Results show localization of Bst protein within the protein storage bodies.

CHART 1

Gonstruction of pUG18pvPpS expression cassette

(a)

HindIII

5'-CCTAAGAAGCTT GGGACTCTGTGGTGTTTTGCTTG-3'

5' Oligomer primer

NcoI

5'-CTGAACCCCATCG TTAGTAGAGTAGTATTGAATATG-3'

3' Oligomer primer

(b)

HindIII NcoI

pvP
500 b.p.

(c)

NcoI

5'-CAATTGGGATGG TTGTATGAACTAAAATGCATGTAG-3'

5' Oligomer primer

HindIII

5'-CAAGTGAAGGTT GTTGAAAGATCCCTGAAGTGTG-3'

3' Oligomer primer

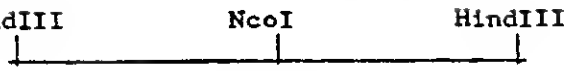
(d)

NcoI HindIII

pvS
1100 b.p.

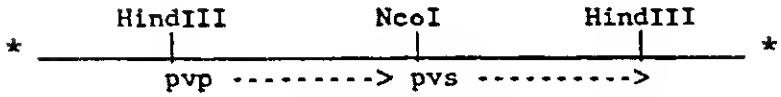
CHART 1 (continued)

(e)



pvPpvS
1600 b.p.

(f)



pUC18pvPpvS

CHART 2

Gonstruction of pUC18pvPpvSPpvTTpvS

(a)

HindIII

5'-GGTAAGAAGCTT GGCACTCTGTGGTCTTTTGGTTC-3'

5' Oligomer primer

NcoI

5'-AATTAATCCATGG CCGGAGTGAAGTGGCAAATG-3'

3' Oligomer primer

(b)

pvPpvSP
580 b.p.

(c)

NcoI

5'-AATGTTCCATGG GCATTTGTGTACTGAATAAGTATG-3'

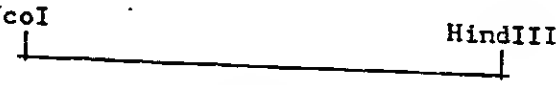
5' Oligomer primer

HindIII

5'-GAACTGAAGGTT GTTGAAAGATCGGTGAAGTGTG-3'

3' Oligomer primer

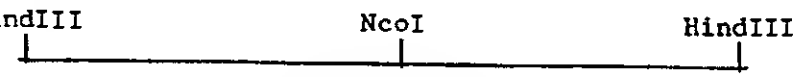
(d)



pvTTpvS
1200 b.p.

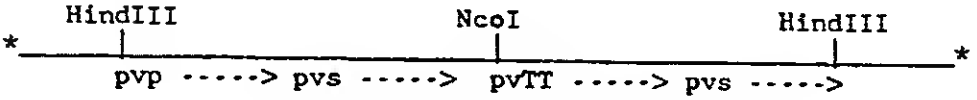
CHART 2 (continued)

(e)



pvPpvSPpvTTpvS

(f)

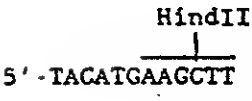


pUC18pvPpvSPpvTTpvS

CHART 3

Construction of pUC18gmPpvS

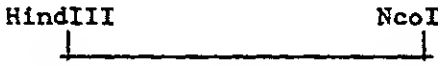
(a)



5' Oligmer primer

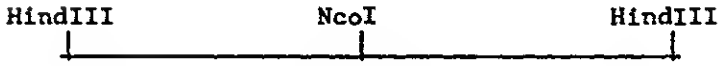
3' Oligomer primer

(b)



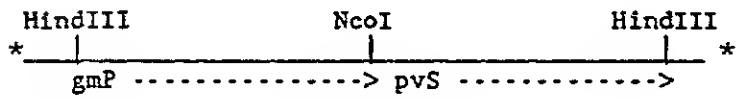
gmP
600 b.p.

(c)



gmPpvS
1700 b.p.

(d)

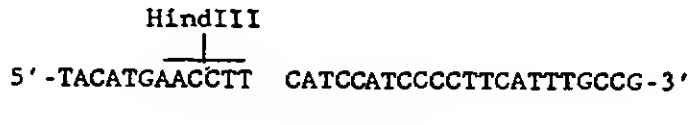


pUC18gmPpvS

CHART 4

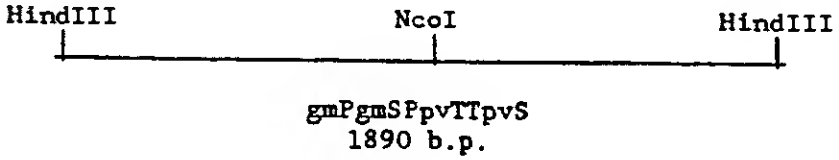
Construction of pUC18gmPgmSPpvTTpvS

(a)



(b)

(c)



(d)

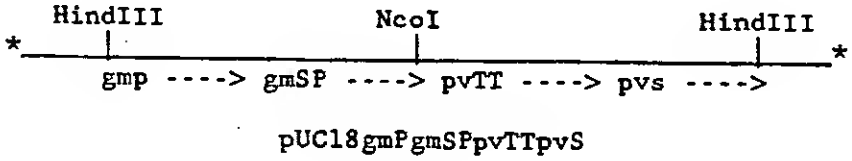


CHART 5

Nucleotide Sequence of pZein15k

```
                GGATGCAAGGCACCTAACAACTAGCACAACATTAGAACAAAATATTGGTGTATATGTGCC
      1         ---------+---------+---------+---------+---------+---------+ 60
10
                TACAATAAGGTGATGAGTCATGGTGATGTGTAAAGAGGCATTACAAAGTTAGCTTCACAA
     61         ---------+---------+---------+---------+---------+---------+ 120
15
                GCGTATGAATTCATTGACAACCGTTGACATGTAAAGTTGATTCATATGTATAAGAAAGCT
    121         ---------+---------+---------+---------+---------+---------+ 180
20
                TAATGATCTATCTGTACATCCAAATCCATGTACTATGTTTCCACgTGATGCAACGCAACA
    181         ---------+---------+---------+---------+---------+---------+ 240
25
                TTGCAAAACCATGGATCATCTATAAATGGCTAGCTCGCACATATGAACTAGTGTCTATGA
    241         ---------+---------+---------+---------+---------+---------+ 300

                TCATCGAATCCAGATCAGCAAAGCGGCAGTGCGTAGAGAGGATCGTCGAACAGAACAGCA
    301         ---------+---------+---------+---------+---------+---------+ 360
30                                                                       M
                TGAAGATGGTCATCGTTCTGGTCGTGTGCCTGGCTCTGTCAGCTGCCAGCGCGTCTGCAA
    361         ---------+---------+---------+---------+---------+---------+ 420
                 K  M  V  I  V  L  V  V  G  L  A  L  S  A  A  S  A  S  A  M
35
                TGCAGATGCCCTGCCCCTGCGCGGGGCTGCAGGGCTTCTACGGCGCTGGCGCCGGCCTGA
    421         ---------+---------+---------+---------+---------+---------+ 480
                 Q  M  P  C  P  C  A  G  L  Q  G  L  Y  G  A  G  A  G  L  T
40
                CGACGATGATGGGCGCGGGGGGCTGTACCCCTACGCGGAGTACCTGAGGCAGCGGCAGT
    481         ---------+---------+---------+---------+---------+---------+ 540
                 T  M  M  G  A  G  G  L  Y  P  Y  A  E  Y  L  R  Q  P  Q  C
45
                GCAGGCCGGTGGGGGCGGGGCCCTACTACGCCGGGTGTGGGCAGACGAGCGCCATGTAGC
    541         ---------+---------+---------+---------+---------+---------+ 600
                 S  P  L  A  A  A  P  Y  Y  A  G  C  G  Q  T  S  A  M  Y  Q
50
                AGCCGCTCCGGCAACAGTGCTGCCAGCAGCAGATGAGGATGATGGACGTGCAGTCCGTCG
    601         ---------+---------+---------+---------+---------+---------+ 660
                 P  L  R  Q  Q  C  C  Q  Q  Q  M  R  M  M  D  V  Q  S  V  A
55
```

CHART 5 (continued)

```
                CGGCAGCAGCTGCAGATGATGATGCAGCTTGAGCGTGCCGCTGCCGGCAGCAGCAGCGTGT
    5    661    ---------+---------+---------+---------+---------+---------+ 720
                 Q  Q  L  Q  M  M  M  Q  L  E  R  A  A  A  A  S  S  S  L  Y

                ACGAGCCAGCTCTGATGCAGCAGCAGCAGCAGCTGCTGGCACTCCAGGGTCTCAACCCCA
   10    721    ---------+---------+---------+---------+---------+---------+ 780
                 E  P  A  L  M  Q  Q  Q  Q  Q  L  L  A  L  Q  G  L  N  P  M

                TGGCCATGATGATGGCGCAGAACATGCGGGCCATGGGTGGACTCTACGAGTACCAGTACC
   15    781    ---------+---------+---------+---------+---------+---------+ 840
                 A  M  M  M  A  Q  N  M  P  A  M  G  G  L  Y  Q  Y  Q  Y  Q

                AGCTGCCCAGCTACCGCACCAACCCCTGTGGCGTCTCCGCTGCCATTGCGCCCTACTACT
   20    841    ---------+---------+---------+---------+---------+---------+ 900
                 L  P  S  Y  R  T  N  P  C  G  V  S  A  A  I  P  P  Y  Y  *

                GATTCATGATATTTGGGAAATCTCCTGTATCCATGCCTCTCTATCTATATATGTAATAAT
   25    901    ---------+---------+---------+---------+---------+---------+ 960

                GCAGTAAGACGACACACATTATCATGTGTGGTATGACCAATAATATATGCATCATAATAA
         961    ---------+---------+---------+---------+---------+---------+1020
   30

                AGTTTTGGTTTTAAAGAATTATCGGACGGTTGATATCTATGATGCTGGATAAATCAAAAC
        1021    ---------+---------+---------+---------+---------+---------+1080

   35

                TTCTCATATAAATTGTAAATATTTGAAAATCTCTATTTAGGCTCCAATGGAGAGGATATG
        1081    ---------+---------+---------+---------+---------+---------+1140

   40           GGTAGAGTAGTATATATGCTTGAAATACTAACAACTAGCAAAGTGCGGGCACGTTGCTAC
        1141    ---------+---------+---------+---------+---------+---------+1200

                ATGCTCATTTATGCTGGAGCATGGAGTATAAAACATAAAGATATATATGTTCTATTGGCT
   45   1201    ---------+---------+---------+---------+---------+---------+1260

                TGGTAAACGTGGATATAGGTTTAAAGCCAACAACTATGGTTCGAATGCTGATTTATATAT
        1261    ---------+---------+---------+---------+---------+---------+1320
   50

                AATCTATAATTTTAGGGTTTTTACTATTTAAATTTTGGGAGTAAGAGGAGAAGGTGCAGG
        1321    ---------+---------+---------+---------+---------+---------+1380

   55
```

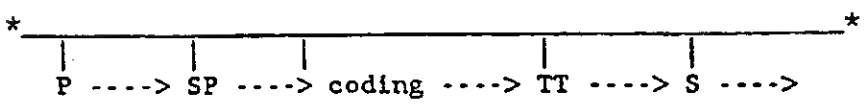
CHART 5 (continued)

```
     GACTTACTCTTGACAACAAATCAATTCGGAAGACACAACAGTAAAAGATAAAATTTCCGA
1381 ---------+---------+---------+---------+---------+---------+1440

     CGGCCAAATGCTTATGCTTCGACGACACCTATCCCTCGGCTCGTAATGGGTCGGAAATAA
1441 ---------+---------+---------+---------+---------+---------+1500

     GGTTATTTCTGACGGCGGCCGTCGGAAATAGCTTATGTTCCGTACGGCCCGAAAGGGCCA
1501 ---------+---------+---------+---------+---------+---------+1560

     TCGATCATTTACCTAAGATAACGATCGTAACGGCCGATGGTGGGTTCCACAGTTCTTATG
1561 ---------+---------+---------+---------+---------+---------+1620

     TCGACGCACTGCTAACCATCGATTATTCGATTCTATGTCTCCGAGGCTTATAGTATTGCG
1621 ---------+---------+---------+---------+---------+---------+1680

     GGCAAACTTATTTGGCATTATGAACTTAACGGAACGATTACATATATATACATATTGATT
1681 ---------+---------+---------+---------+---------+---------+1740

     CATATAC
1741 ------- 1747
```

CHART 6

Construction of pUC18zmPzmS

(a)



(b)

SalI

5' -ATAGCTGTCCAC GATCCAAGGCATCTAACAACT-3'

5' oligomer primer

NcoI

5'-CTCCTACCATGG TCTTCTCTTCCACCATCCTCTCTA-3'

3' oligomer primer

(c)

zmP
350 b.p.

(d)

NcoI

5'-CATAGCCATCC TATTTCCGAAATCTCCTTC-3'

5' oligomer primer

SalI

5'-TAGACTGTCGAC AAGAACTGTGCAACCCACCATCC-3'

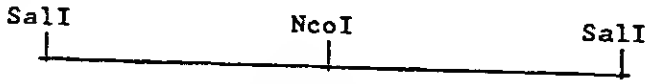
3' oligomer primer

CHART 6 (continuation)

(e)

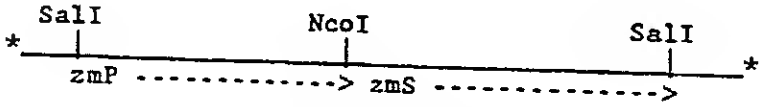
zmS
700 b.p.

(f)



zmPzmS
1050 b.p.

(g)



pUC18zmPzmS

CHART 7

Construction of pUC18zmPzmSPzmTTzmS

(a)

SalI

5'-ATAGCTGTCGAG GATCCAAGGCATCTAACAACT-3'

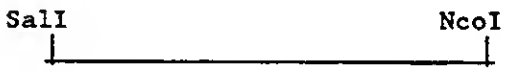
5' oligomer primer

NcoI

5'-AATTACCCATGG GCAGGGCATCTGCATTGCAGAG-3'

3' oligomer primer

(b)



zmPzmSP
400 b.p.

(c)

NcoI

5'-AATCATCCATGG GCCTACTACTGATTCATGATATTTG-3'

5' oligomer primer

SalI

5'-TAGACTGTCGAC AAGAACTGTGGAACCCACCATCG-3'

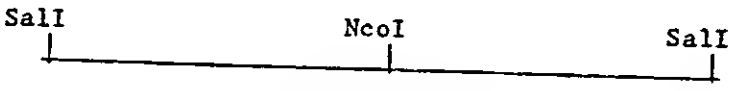
3' oligomer primer

(d)

zmTTzmS
720 b.p.

CHART 7 (continuation)

(e)



zmPzmSPzmTTzmS
1120 b.p.

(f)

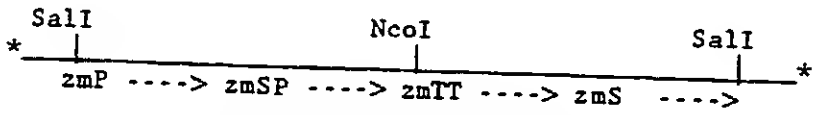


CHART 8

Nucleotide Sequence of Brazil Nut
Sulfur-Rich Seed Storage Protein

```
                TCCACCACTGCTCTGTATCACATATACCAACCCACGCTTATACCCACAATCACCATGCCG
10      1       ---------+---------+---------+---------+---------+---------+  60
                                                                  M   A

                AACATTTCACTTCCCCCACCAGGCCTCCTTGTCCTCATCGCCCTCGCCCACCCCACCCCC
15      61      ---------+---------+---------+---------+---------+---------+ 120
                K   I   S   V   A   A   A   A   L   L   V   L   M   A   L   C   H   A   T   A

                TTCCCCCCCACCCTCACCACCAGACTCCTGGACCACCACAACCACGACGAGTCTGGCGAG
20      121     ---------+---------+---------+---------+---------+---------+ 180
                F   R   A   T   V   T   T   T   V   V   E   E   E   N   Q   E   E   G   R   E

                CAGATCCAGAGACACCACATGGTCAGCCACTGCGGGATGTACATGAGACAGCACATCCAC
25      181     ---------+---------+---------+---------+---------+---------+ 240
                Q   M   Q   R   Q   Q   M   L   S   H   C   R   M   Y   M   R   Q   Q   M   E

                CAGACCCCCTACCAGACCATGCCCACGCCCCCAATCCACCCCCACATGACCCACTCCTCC
30      241     ---------+---------+---------+---------+---------+---------+ 300
                E   S   P   Y   Q   T   M   P   R   R   G   M   E   P   H   M   S   E   G   C

                GACCACCTCGAGCGGATGGACCACACCTGCAGATCCGAAGGCTTAACGATCATGATGATG
35      301     ---------+---------+---------+---------+---------+---------+ 360
                E   Q   L   E   G   M   D   E   S   C   R   C   E   G   L   R   M   M   M   M

                AGGATGCAACAGGAGGAGATGCAACCCCGAGGGGAGCAGATGCGAAGGATGATGACCCTC
40      361     ---------+---------+---------+---------+---------+---------+ 420
                R   M   Q   Q   E   E   M   Q   P   R   C   E   Q   M   R   R   M   M   R   L

                GGCCAGAATATCCCTTCCCGCTGCAACCTCAGTCCCATGAGATGCCCCATGCCTCCCTCC
45      421     ---------+---------+---------+---------+---------+---------+ 480
                A   E   N   I   P   S   R   C   N   L   S   P   M   R   C   P   M   C   G   S

                ATTCCGCCGTTCTCAATCTGCCACTAGCCAGTGCTCTAAATCTTAATAACCCTCTCACAA
50      481     ---------+---------+---------+---------+---------+---------+ 540
                I   A   G   F   *
```

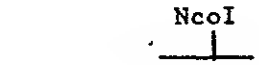
CHART 8 (continuation)

```
5          ACTAGGTGTTTGTTGGCTTTTGGCCGGAGACTAGGGTGTGGGATTAATAATAATAGCACA
      541  ---------+---------+---------+---------+---------+---------+ 600

            CTATGGTGTGTTCTCAGCTTC
10    601  ---------+---------+-  621
```

CHART 9

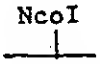
Construction of Expression Cassettes Using the Brazil Nut Sulfur-Rich Seed Storage Protein Gene (BZ)

(a)



5'-AATTCCACCATGG CGAAGATTTCAGTTG-3'

5' oligomer primer



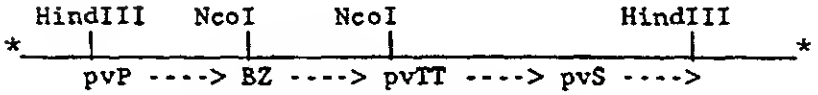
5'-AATTCAACCATGG CATGGGGGCATCTCATGGGACTG-3'

3' oligomer primer

(b)

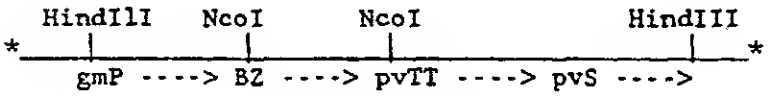
BZ
420 b.p.

(c)(i)



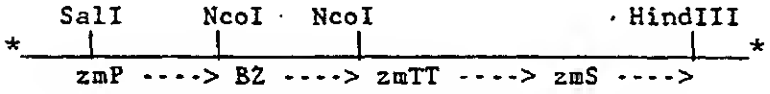
pUC18pvPBZpvTTpvS

(c)(ii)



pUC18gmPBZpvTTpvS

(c)(iii)



pUC18zmPBZzmTTzmS

CHART 10

(a) Nucleotide Sequence of a Synthetic 9 kD subunit of a Brazil Nut Sulfur-Rich Seed Storage Protein (BZ-1)

```
      ATGGAGCCGGACATGAGCGAGTGCTCCGAGCAGCTGGAGGGGATGGAGGAGAGCTGCAGA
   1  ---------+---------+---------+---------+---------+---------+  60
      M  E  P  H  M  S  E  G  G  E  Q  L  E  G  M  D  E  S  C  R

      TCCGAAGGCTTAAGGATCATGATGATGAGGATGCAACAGGAGCAGATGCAACCGCGAGGG
  61  ---------+---------+---------+---------+---------+---------+ 120
      C  E  G  L  R  M  M  M  M  R  M  Q  Q  E  E  M  Q  P  R  G

      GAGCAGATGGGAACGATGATGAGG
 121  ---------+---------+----  144
      E  Q  M  R  R  M  M  R
```

(b) Nucleotide Sequence of a Synthetic Brazil Nut Sulfur-Rich Seed StorageProtein Complete 9 kD Gene (BZ-2)

```
      GCCAGGCGGGGAATGGAGCCGGAGATGAGGGAGTGGTGGGAGGAGGTGGAGGGGATGGAC
   1  ---------+---------+---------+---------+---------+---------+  60
      P  R  R  C  M  E  P  H  M  S  E  G  C  E  Q  L  E  G  M  D

      GAGAGCTGCAGATGGGAAGGCTTAAGCATGATGATGATGAGGATGCAAGACCAGGACATG
  61  ---------+---------+---------+---------+---------+---------+ 120
      E  S  G  R  G  E  G  L  R  M  M  M  M  R  M  Q  Q  E  E  M

      CAAGCCCCAGGGGACCACATGCGAAGGATGATGAGGCTGGCCGAGAATATGCCTTCCCGC
 121  ---------+---------+---------+---------+---------+---------+ 180
      Q  P  R  G  E  Q  M  R  R  M  M  R  L  A  E  N  I  P  S  R

      TCCAACCTCACTCCCATGAGATGCCCCATC
 181  ---------+---------+---------+  210
      C  N  L  S  P  M  R  C  P  M
```

CHART 10 (continued)

(c) Nucleotide Sequence of a Synthetic Brazil Nut 12 kD Subunit of the Sulfur-Rich Seed Storage Protein Gene (BZ-3)

```
          CAGCAGATGCTCAGCCACTGCCGGATGTACATGAGACAGCAGATGGAGGAGAGCCCGTAC
     1    ---------+---------+---------+---------+---------+---------+   60
          Q  Q  M   L  S  H  C  R  M  Y  M  R  Q  Q  M  E  E  S  P  Y

          CAGACCATGCCCAGGCGGGGAATGGAGCCGCACATGAGCGAGTGCTGCGAGCAGCTGGAG
    61    ---------+---------+---------+---------+---------+---------+  120
          Q  T  M   P  R  R  G  M  E  P  H  M  S  E  C  C  E  Q  L  E

          GGGATGGACGAGAGCTGCAGATGCGAAGGCTTAAGGATGATGATGATGAGGATGCAACAG
   121    ---------+---------+---------+---------+---------+---------+  180
          G  M  D  E  S  C   R  C  E  G  L  R  M  M  M  M  R  M  Q  Q

          GAGGAGATGCAACCCCGAGGGGAGCAGATGCGAAGGATGATGAGGCTCGCCGAGAATATC
   181    ---------+---------+---------+---------+---------+---------+  240
          E  E  M   Q  P  R  G  E  Q  M  R  R  M  M  R  L  A  E  N  I

          CCTTCCCGCTGCAACCTCAGTCCCATGAGATGCCCCATG
   241    ---------+---------+---------+---------   279
          P  S  R  C  N  L  S  P  M  R  C  P  M
```

CHART 11

(a)

NcoI

5'-AATTATCCATGG ATGCAGCCCCACATGAGCGAGTG-3'

5' oligomer primer

NcoI

5'-AATTCTCCATGC CATCATCCTTCCCATCTGCTC-3'

3' oligomer primer

(b)

NcoI

5'-AATTCACCATGC CCAGGCGGGGAATCCAGCCGC-3'

5' oligomer primer

NcoI

5'-AATTCACCATGG CATGGGGCATCTCATGCCACTG-3'

3' oligomer primer

(c)

NcoI

5'-ATTCACCATCC CACCACATCCTCACCCACTGCC-3'

5' oligomer primer

NcoI

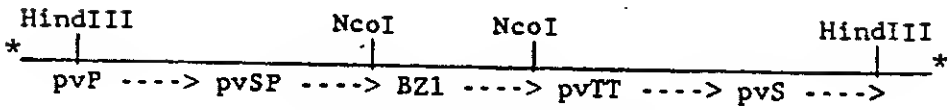
5'-AATTCACCATGG CATGGCGCATCTCATGGGACTG-3'

3' oligomer primer

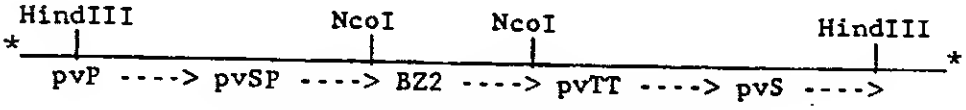
CHART 12

Phaseolin Expression Cassettes Using
Brazil Nut Cene Fragments

(a)



(b)



(c)

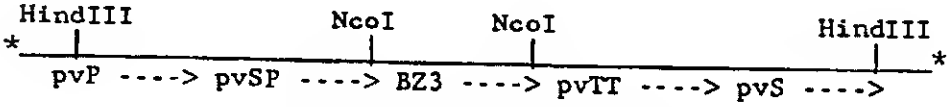
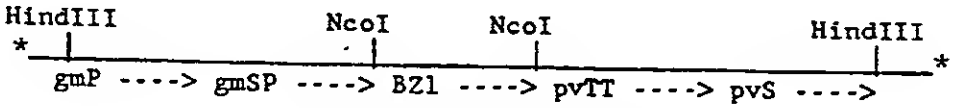


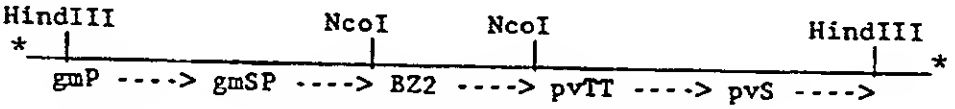
CHART 13

Glycine Expression Cassettes Using
Brazil Nut Gene Fragments

(a)



(b)



(c)

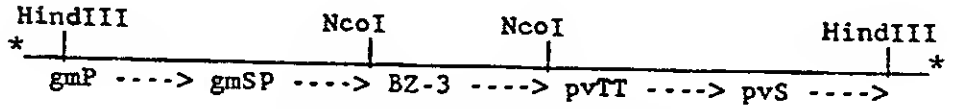
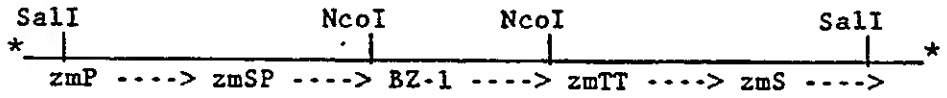


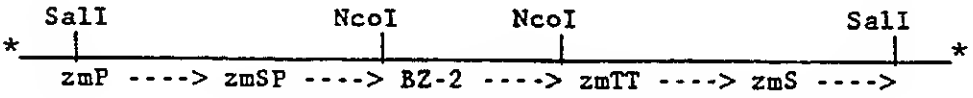
CHART 14

Zein Expression Cassettes Using
Brazil Nut Gene Fragments

(a)



(b)



(c)

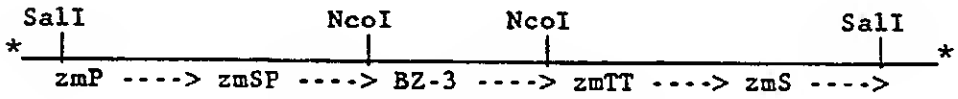
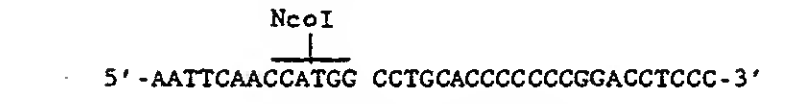


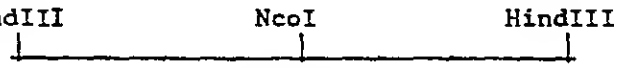
CHART 15

Phaseolin Expression Cassettes Using BST Gene

(a)



(b)



(c)

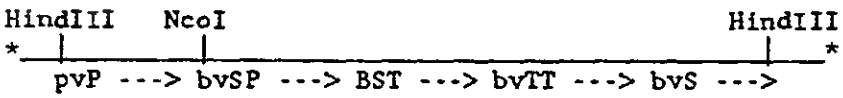
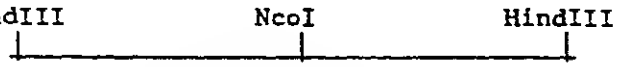


CHART 16

Glycine Expression Cassettes Using BST Gene

(a)



(b)

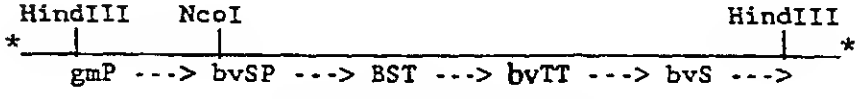
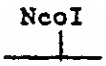


CHART 17

Phaseolin Expression Cassettes Using BST Gene

(a)



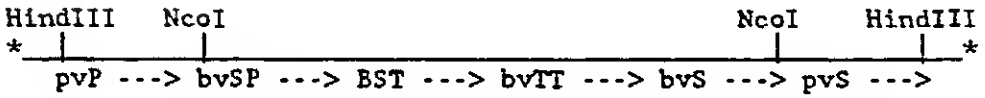
5'-AAATTACCATGG GCATCCCCACCATCCCTCCTATTG-3'

3' oligomer primer

(b)

bvSPBSTbvTTbvS

(c)



pUC18pvPbvSPBSTbvTTbvSpvS

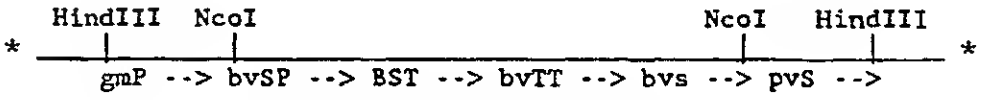
CHART 18

Clycine Expression Cassettes Using BST Gene

(a)

bvSPBsTbvTTbvS

(b)



pUC18gmPbvSPBSTbvTTbvSpvS

CHART 19

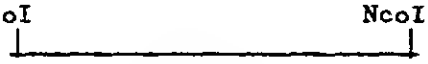
Phaseolin Expression Cassettes Using BST Gene

(a)

5'-AATTCAGGATGG GGGGGAGGGGGAAAGAACAGATGG-3'

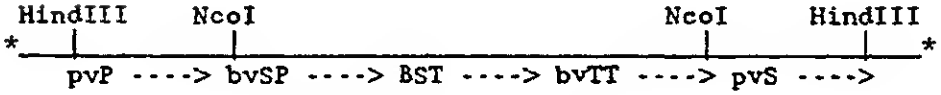
3' oligomer primer

(b)



pvSPBSTbvTT

(c)



pUC18pvPbvSPBSTbvTTpvS

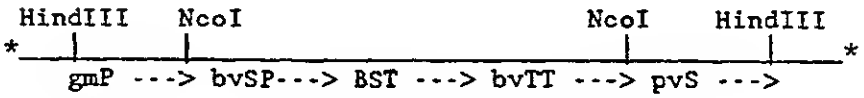
CHART 20

Glycine Expression Cassettes Using BSt Gene

(a)

pvSPBSTbvTT

(b)

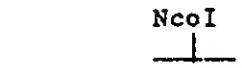


pUC18gmPbvSPBSTbvTTpvS

CHART 21

Phaseolin Expression Cassettes Using BST Gene

(a)



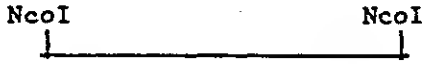
5'-AATCATCCATGG TTCCCAGCCATGTCCTTGTCCGGCC-3'

5' oligomer primer

5'-AATCACCATGG GGGGGAGGGGCAAACAACAGATGG-3'

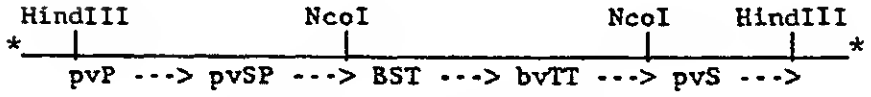
3' oligomer primer

(b)



BSTbvTT

(c)

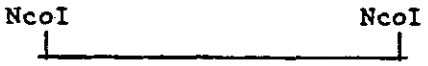


pUC18pvPpvSPBSTbvTTpvS

CHART 22

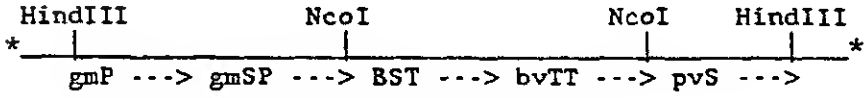
Clycine Expression Cassettes Using BST Gene

(a)



BSTbvTT

(b)



pUC18gmPgmSPBSTbvTTpvS

CLAIMS

We claim:

1. A seed specific expression cassette having gene regulatory elements comprising:

a) a promoter selected from the group consisting of phaseolin promoter, α'-subunit of β-conglycinin promoter, and β-zein promoter;

b) a translation initiation signal selected from the group consisting of phaseolin translation initiation signal, α'-subunit of β-conglycinin translation initiation signal, β-zein translation initiation signal, and animal gene translation initiation signal;

c) a gene selected from the group consisting of an animal gene and a gene from brazil nut 2S storage protein;

d) a poly-A signal selected from the group consisting of a phaseolin poly-A signal and an animal gene poly-A signal;

wherein the promoter is upstream and operably linked to the initiation signal, the initiation signal is upstream and operably linked to the gene, and the gene is upstream and operably linked to the poly(A) signal; and provided that when the gene regulatory element is amplified from genomic or synthetic DNA the regulatory region is flanked by a restriction enzyme recognition site selected from the group consisting of NcoI, HindIII, and SalI; and further provided when the translation initiation signal is from phaseolin the promoter is from phaseolin, and when the translation initiation signal is from the α'-subunit of β-conglycinin the promoter is from the α'-subunit of β-conglycinin, and when the gene is from brazil nut the promoter and translation initiation signal are derived from the same genome.

2. A seed specific expression cassette according to Claim 1 wherein the promoter is from phaseolin, the remaining regions are from bovine somatotropin, and further comprising a SV40 promoter region upstream and operably linked to the phaseolin region.

3. A seed specific expression cassette according to Claim 1 wherein the gene is an animal gene from bovine somatotropin and includes the translation initiation signal.

4. A seed specific expression cassette according to Claim 3 wherein

the promoter is from phaseolin, the remaining elements are from bovine somatotropin, the regions linking both the promoter and the poly-A elements to the plasmid each code for HindIII, and the regions linking the promoter to the translation initiation signal codes for NcoI.

5. A seed specific expression cassette according to Claim 3 wherein the promoter is from the α'-subunit of β-conglycinin, the remaining elements are from bovine somatotropin, the regions linking both the promoter and the poly-A elements to the plasmid each code for HindIII, and the region linking the promoter to the translation initiation signal codes for NcoI.

6. A seed specific expression cassette according to Claim 3 wherein the promoter and the poly-A signal are from phaseolin, the remaining elements are from bovine somatotropin, the region linking both the promoter and the poly-A elements to the plasmid each code for HindIII, and the region linking the promoter to the translation initiation signal and the gene to the poly-A signal code for NcoI.

7. A seed specific expression cassette according to Claim 3 wherein the promoter and the poly-A signal are from the α'-subunit of β-conglycinin, the remaining elements are from bovine somatotropin, the region linking both the promoter and the poly-A elements to the plasmid each code for HindIII, and the region linking the promoter to the translation initiation signal and the gene to the poly-A signal code for NcoI.

8. A seed specific expression cassette according to Claim 3 wherein the promoter, translation initiation signal, and poly-A signal are from phaseolin, the region linking both the promoter and the poly-A elements to the plasmid each code for HindIII, and the region linking the translation initiation signal to the gene and the gene to the poly-A signal code for NcoI.

9. A seed specific expression cassette according to Claim 3 wherein the promoter is from phaseolin, the translation initiation signal and

gene are from bovine somatotropin, and having a first poly-A signal from bovine somatotropin upstream and operably linked to a second poly-A signal from phaseolin, the region linking both the promoter and the second poly-A elements to the plasmid each coding for HindIII, and the region linking the promoter to the translation initiation signal and the first poly-A signal to the second poly-A signal coding for NcoI.

10. A seed specific expression cassette according to Claim 3 wherein the promoter is from the α'-subunit of β-conglycinin, the translation initiation signal and gene are from bovine somatotropin, and having a first poly-A signal from bovine somatotropin upstream and operably linked to a second poly-A signal from phaseolin, the region linking both the promoter and the second poly-A element to the plasmid each coding for HindIII, and the region linking the promoter to the translation initiation signal and the first poly-A signal to the second poly-A signal coding for NcoI.

11. A seed specific expression cassette according to Claim 1 comprising:

a) a promoter selected from the group consisting of phaseolin promoter, α'-subunit of β-conglycinin promoter, and β-zein 15K promoter.

b) a translation initiation signal selected from the group consisting of phaseolin translation initiation signal, α'-subunit of β-conglycinin translation initiation signal, and β-zein 15K translation initiation signal;

c) a gene selected from the group consisting of brazil nut gene and brazil nut gene subunit;

d) a translation termination signal selected from the group consisting of phaseolin translation termination signal and β-zein 15K translation termination signal;

e) a poly-A signal selected from the group consisting of phaseolin poly-A signal and β-zein 15K poly-A signal;

wherein the promoter is upstream and operably linked to the initiation signal, the initiation signal is upstream and operably linked to the gene, the gene is upstream and operably linked to the translation

termination signal, and the translation termination signal is upstream and operably linked to the poly-A signal, and further provided that the gene regulatory element is flanked by a restriction enzyme recognition site selected from the gr up consisting of NcoI, HindIII, and SalI; and further provided that the promoter and translation initiation signals are derived from the same genome.

12. A seed specific expression cassette according to Claim 11 wherein the gene is BZ.

13. A seed specific expression cassette according to Claim 12 wherein the promoter, translation termination signal, and poly-A signal are from phaseolin, the region li king both the promoter and poly-A elements to the plasmid each code for HindIII, and the region linking the promoter to the gene and the gene to the termination signal code for NcoI.

14. A seed specific expression cassette according to Claim 12 wherein the promoter is from the α'-subunit of β-conglycinin, the translation termination signal and poly-A signal are from phaseolin, the region linking both the promoter and poly-A elements to the plasmid each code for HindIII, and the region linking the promoter to the gene and the gene to the termination signal code for NcoI.

15. A seed specific expression cassette according to Claim 12 wherein the promoter, translation termination signal, and poly-A signal are from β-zein 15K, the region linking both the promoter and poly-A elements to the plasmid each code for SalI, and the regions linking the promoter to the gene and the gene to the termination signal code for NcoI.

16. A seed specific expression cassette according to Claim 11 wherein the gene is selected from the group consisting of BZ-1, BZ-2, and BZ-3.

17. A seed specific expression cassette according to Claim 16 wherein the remaining elements are from phaseolin, the region linking

both the promoter and poly-A elements to the plasmid each code for HindIII, and the region linking the translation initiation signal to the gene and the gene to the termination signal code for NcoI.

18. A seed specific expression cassette according to Claim 16 wherein the remaining elements are from β-zein 15K, the region linking both the promoter and poly-A elements to the plasmid each code for SalI, and the region linking the translation initiation signal to the gene and the gene to the termination signal code for NcoI.

19. A cell comprising a seed specific expression cassette as claimed in Claim 1.

20. A seed or plant comprising cells as claimed in Claim 19.

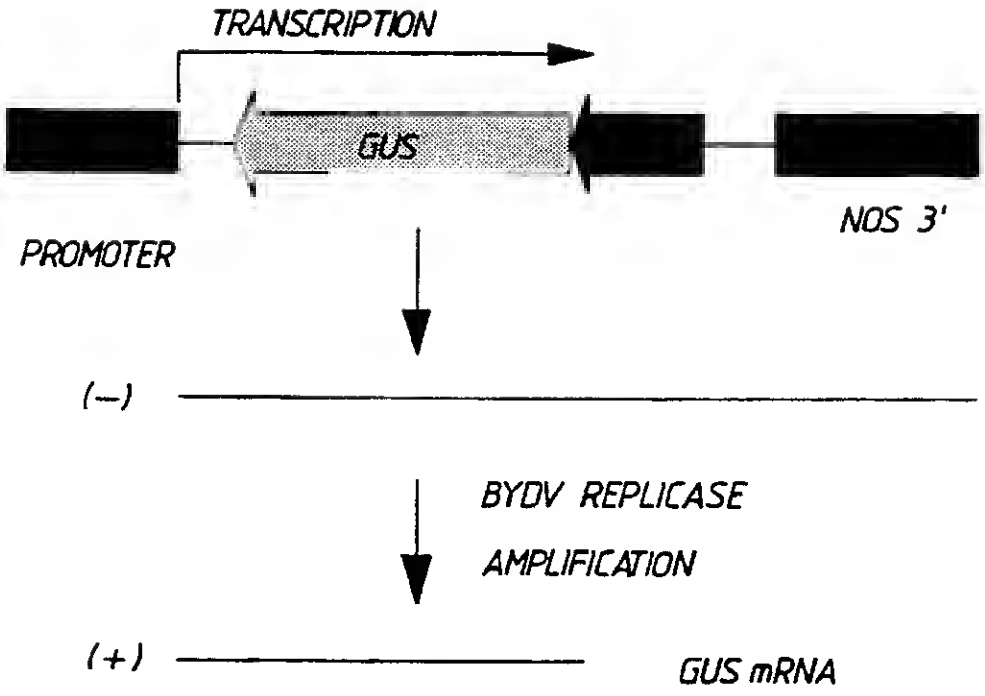
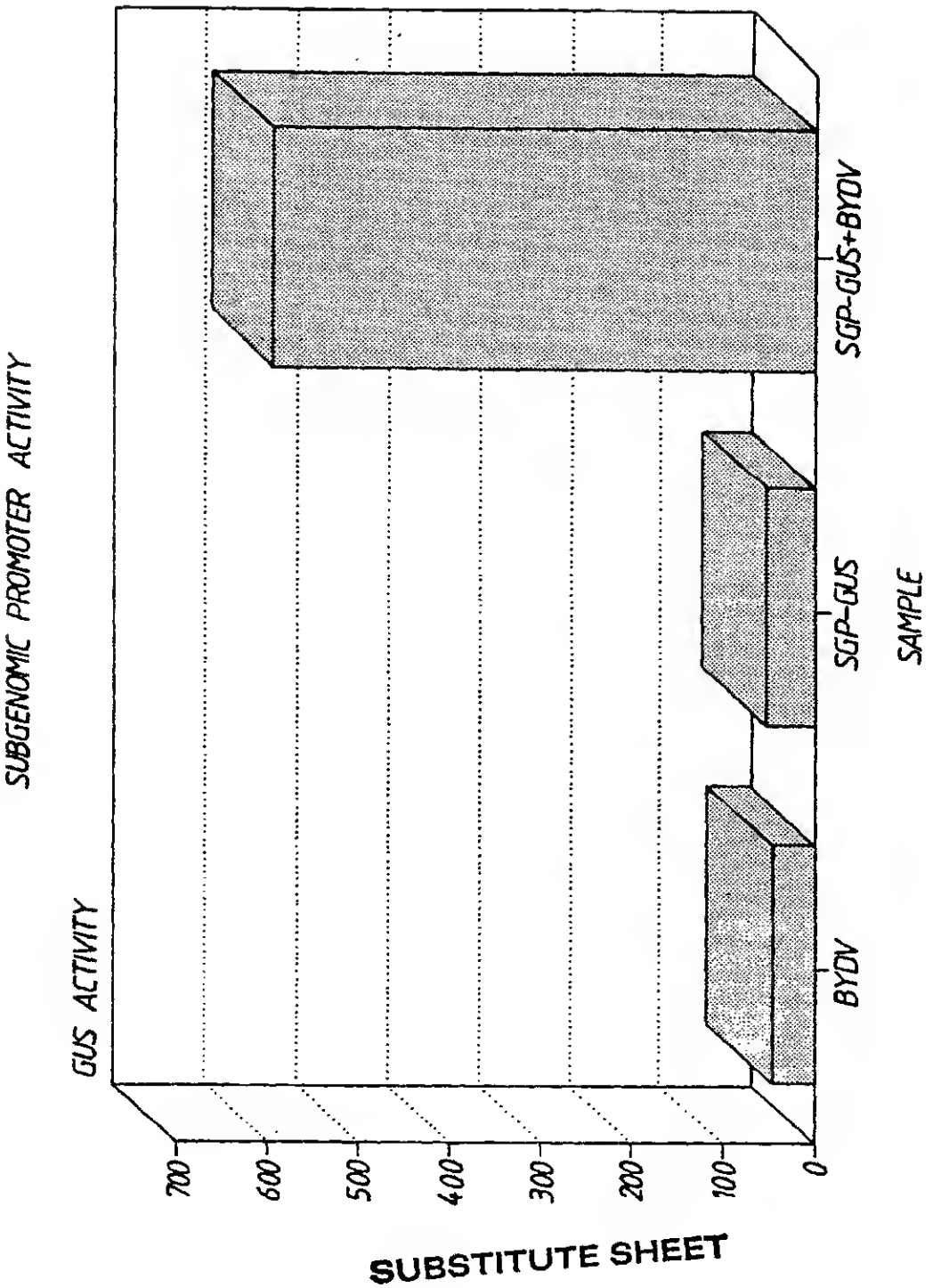


Fig.1.

Fig.2.



INTERNATIONAL SEARCH REPORT

International Application No PCT/US 91/00887

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 12 N 15/82, C 12 N 5/10, A 01 H 5/00

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 12 N, A 01 H

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	EP, A, 0295959 (PLANT CELL RESEARCH INST.) 21 December 1988 see page 5, column 8, lines 52-61; page 9, column 15, line 20 - column 16, line 20 --	1,11-13, 16,17,19, 20
A	WO, A, 8402913 (MONSANTO) 2 August 1984 see pages 41-44 --	2-10,19, 20
A	WO, A, 8700865 (CALGENE, INC.) 12 February 1987 see page 4, lines 16-17; page 9, line 32 - page 10, line 6 --	2-10,19, 20
A	The EMBO Journal, vol. 7, no. 2, 1988, IRL Press Ltd, (Oxford, GB), Z.-L. Chen et al.: "A DNA sequence element that confers seed-specific enhancement to a constitutive promoter", pages 297-302 see the whole article --	1

./.

[10] Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
5th July 1991	30. 08. 91
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	M. PEIS M. Peis

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, ** with indication, where appropriate, of the relevant passages	Relevant to Claim No.
A	WO, A, 9001551 (ROGERS) 22 February 1990 see page 13, lines 1,2 --	2-10,19,20
A	Plant Molecular Biology, vol. 13, November 1989, Kluwer Academic Publishers, (BE), S.B. Altenbach et al.: "Enhancement of the methionine content of seed proteins by the expression of a chimeric gene encoding a methionine-rich protein in transgenic plants", pages 513-522 see the whole article --	1,11-13,16, 17,19,20
A	Journal of Cellular Biochemistry, suppl. 12C, 1988, Alan R. Liss, Inc., (New York, US), see abstract L300 S.B. Altenbach et al.: "Processing of a Brazil nut sulfur-rich seed protein in transgenic plants", page 177 --	1,11-13,16, 17,19,20
A	EP, A, 0319353 (PGS) 7 June 1989 see the whole document cited in the application ----------	2-10,19,20

Form PCT/ISA 210(extra sheet) (January 1985)

ANNEX TO THE INTERNATIONAL SEARCH REPORT
ON INTERNATIONAL PATENT APPLICATION NO. US 9100887
SA 46073

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.
The members are as contained in the European Patent Office EDP file on 26/08/91
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0295959	21-12-88	AU-A- 1810088	22-12-88
		JP-A- 1091787	11-04-89
WO-A- 8402913	02-08-84	EP-A,B 0131623	23-01-85
		JP-T- 60500796	30-05-85
WO-A- 8700865	12-02-87	US-A- 4956282	11-09-90
		AU-B- 603063	08-11-90
		AU-A- 6196686	05-03-87
		EP-A- 0233915	02-09-87
		JP-T- 63500425	18-02-88
WO-A- 9001551	22-02-90	AU-A- 4037289	05-03-90
		EP-A- 0428572	29-05-91
EP-A- 0319353	07-06-89	AU-A- 2811889	23-05-89
		WO-A- 8903887	05-05-89
		EP-A- 0318341	31-05-89
		JP-T- 2501802	21-06-90

EPO FORM P0479

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82